The Effect of Education on Income Inequality

Matthew Curtis

March 22, 2021

Econ 3161: Dr. Shatakshee Dhongde

Abstract

This paper shines a light on the relationship between education attained and income inequality. All of our data is recent and comes from the Human Development Data Center. The income inequality was measured by the Gini Coefficient, a renowned statistic used to measure the income inequality of a country. A simple linear regression was evaluated by using country's average number of years of education completed and the country's Gini coefficient. After reviewing multiple studies attributed to this relationship, it was apparent that there are multiple other factors contributing to this relationship. Multiple regression models were calculated with secondary variables such as government expenditure and median age. Understanding this relationship will help countries decrease income inequality.

Introduction

As the gap in wage rates continues to increase across various countries, it is important to assess the reasoning behind these changes. Governments are intrigued by the increasing gap, as they hope to support nationwide programs to decrease this gap. There are many effects on wage rates, ranging from experience in the field, skill, age, government expenditure, gender, and education. Though, it is becoming increasingly more apparent how significant education is when discussing wage rates. Many studies have been conducted to assess the difference in income by household when comparing the level/amount of education attained by the head of the household. There's a noticeable link between amount of education and income earned, but it is intriguing to find how significant the relationship truly is, and it is interesting to see how big of a factor these other variables (age, government expenditure, etc.) are when discussing this relationship.

Comparing generation to generation, the difference in amount of people receiving higher levels of education is very apparent. From personal experience, in the US kids are able to realize the difference in prioritization of education when comparing their lives to their parents. There has been a growing amount of emphasis placed on attaining higher levels of education almost across the entire world. Many believe these extra years of education will result in finding a better-paying, more reliable job. A person's societal status is also commonly attributed with the level of education they have received. The believed benefits of education go beyond societal status or income, as education has positive national effects of reduced crime rate, better public health, and an improved approach to parenthood.

This paper aims to aid in knowing the true impact of education on a country's income inequality. Specifically, I desire to know the impact of education on decreasing the wage gap, and I want to have a better idea of what actions need to be taken to successfully decrease this gap. I believe a negative correlation will be examined between level of education attained and a country's income inequality across almost every country studied. If a strong correlation is found then the economy and workforce can be massively improved by focusing government policies on incentivizing individuals to attain higher levels of education. Today it is implicitly agreed that countries should put forth good deals of effort into making education accessible and important to everyone, but are these efforts the most efficient ways to allow households below the poverty line to positively change their income situation? Some argue that this is actually not the most efficient way to decrease the wage gap nationally.

Literature Review

John Jerrim and Lindsey Macmillan (2015) aim to find the impact of education on income inequality, the financial returns on investing in education, and the effect of education on labor-market earnings. Their paper was created with hopes of filling the gap of little evidence connecting education to "social origin and destination." They also shined a light on the Great Gatsby Curve, the relationship between income inequality and intergenerational mobility. This relationship is showed to readers by plotting the Gini Coefficient, a measure of income inequality, against intergenerational income elasticity, a measure of social mobility. The graph represents the association between less social mobility and higher income inequality within a country. They realize how changing social status within a family can be incredibly difficult, so they knew this was a massive factor when considering the relationship between education and income inequality. Jerrim and Macmillan explain how income inequality is becoming worse and more noticeable in developed countries. After explaining this massive influence, data was collected from 24 different developed countries to examine the financial returns on education. Financial returns on education are defined as the earnings of the graduates, and the earnings of university graduates was compared to the earnings of high school graduates. Earnings were all converted into a reporting period that was consistent across all data points. Upon testing and evaluating data, they have determined the economic return on education is high. In certain countries the university graduates were seen to have a wage return of about 60 percent more than high school graduates. There is a noticeable link between educational inequality and income inequality.

Juan Yang and Man Gao (2017) similarly investigate the impact of education expansion on wage inequality. They realize the importance of their findings, as the conclusions could heavily impact the decisions policymakers face when trying to reduce income inequality. The authors realize that education has been seen as a reliable way for individuals to earn more income, and they aim to truly see if this theory holds true. There has been an upward trend in the number of students enrolled in higher education across many different countries. The authors point out the difference of income between a high school graduate versus a college graduate, and how the market return to college graduates has been continuously increasing. Yang and Gao introduce the idea to readers that education can definitely increase an individual's income, but they are interested in the idea of if government intervention in education expansion will have a beneficial impact on reducing income inequality. After comparing data and analyzing the rates of return to education for college graduates versus high school graduates, it is noted that expanding education will decrease the income gaps by the structure effect. On the other hand, the price effect of education expansion is positive, and its effect is much more significant than the structure effect. This significant effect leads us to believe that expanding education will actually increase the income inequality.

Jana Turcinkova and Jana Stavkova (2009) aim to prove that attaining a higher level of education is no guarantee of lower risk of poverty. They do admit that amount of education attained is linked with income inequality and risk of poverty, but they intend to examine how powerful that relationship is. Turcinkova and Stavkova collected data on the head of household's level of education and income situation in the Czech Republic from 2005-2009. Their goal was to specifically analyze the income vulnerable households by surveying their attained level of education. They also shined a light on other variables such as household structure, economic activity, and social situation. After analyzing trends from 2005-2009 it was clear that the most poverty-vulnerable group was households that had primary education or no education at all. The number of vulnerable households in this category was a little more than double the average of poverty-vulnerable households across the country.

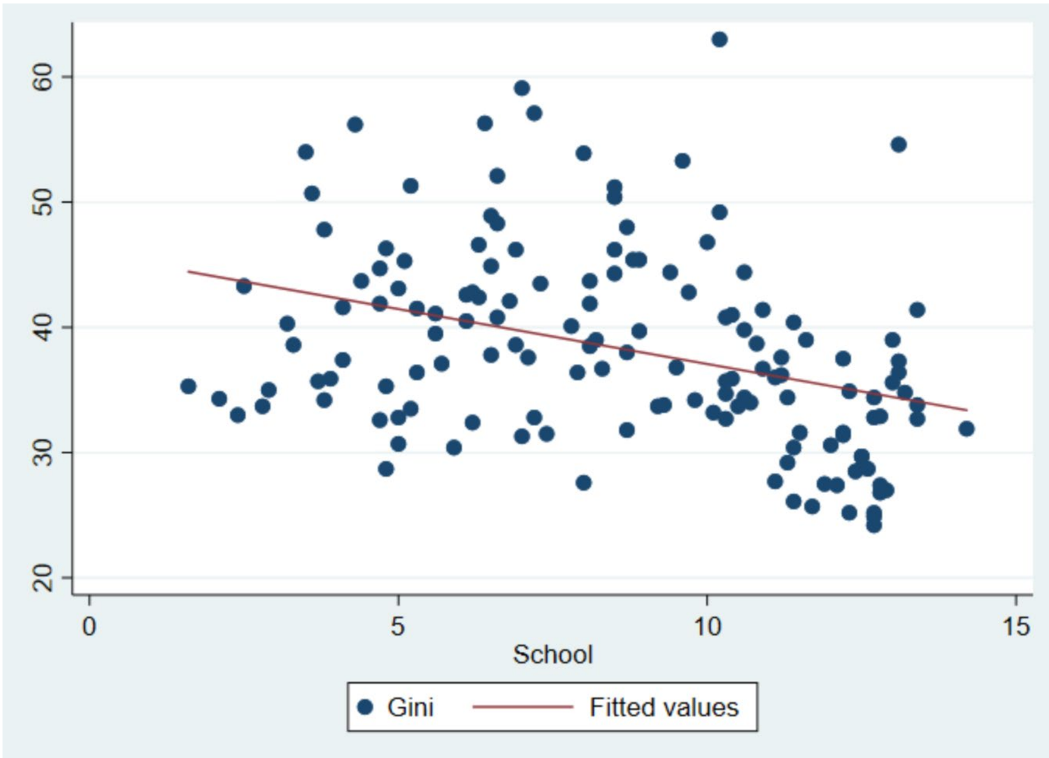
Abdul Abdullah (2013) of Deakin University conducted a study to find the effects of education on the distribution of income in the country of Africa. Abdullah also aims to find the other major factors that would impact the relationship between education and income. It has been evident that government intervention has a large impact on this relationship, as educational priorities have been set in various countries. Abdullah discusses the expansion of education from the government may not benefit the poor as the programs were intended to do; those of higher income may be able to take advantage of the better educational opportunities compared to those who may struggle with not having sufficient resources to attend school. This leads us to believe that government subsidizing education may disproportionately impact the wealthy which would increase income inequality within the country. Aside from government intervention, Abdullah was able to conclude education does reduce the gap between the rich and the poor. Education has a more significant impact on increasing the income of the poor than decreasing the income of the rich. The results suggest that completing secondary schooling has a more significant impact compared to the completion of primary school. Knowing this, the author concludes that government subsidization towards secondary schooling could be the most effective action. He believes the overall expansion of education would not be as beneficial as specifically creating policies in attempt to increase individuals completing secondary schooling. Abdullah realizes the short-comings of his studies though, as he points out that the results do not provide information regarding the cost-benefit analysis of education.

As seen above, there is a lot of research analyzing the relationship between education attained and a country's income inequality. This research is different because it takes into account the previous studies above, and it includes the variables that are most influential regarding this relationship. After reviewing these studies, I have determined that the secondary explanatory variables of government expenditure and median age of the country will be beneficial to use when finding the strength of this relationship. Many of these studies above focused on a very specific data set, or a specific country, but I will be using data from every country available. This will give us the best understanding of the correlation between education and income inequality across the entire world.

Data

In an effort to characterize the relationship between income inequality and amount of education attained within a country, cross sectional data was analyzed. The dependent variable was defined by the Gini Coefficient, a measure of a country's income inequality. The Gini Coefficient is measured on a unitless scale, ranging from 0 to 100, 0 representing perfect income equality, while 100 represents perfect income inequality within a country. After reviewing a variety of literature regarding the relationship between income inequality and education, it is safe to say that the Gini Coefficient is widely regarded as the best way to measure a country's income inequality. The main explanatory variable is the average amount of education attained. This variable was measured by surveying people ages 25 and older, in attempt to exclude the individuals currently receiving education. To keep consistency across analysis, the average number of years of education was converted "from education attainment levels using official durations of each level." This is in response to the fact that the definition of "years" of education varies from country to country. Initially, it was desired to find data about the average level of education completed from country to country, but it was realized that this was very inefficient and inconsistent due to the wide variety of educational structures from country to country. As seen in Figure 1, an initial scatterplot of the dependent variable vs. the main explanatory variable shows a slight negative correlation. Intuitively, this is expected. The higher education attained is expected to correlate with a smaller Gini coefficient for the country, meaning there is less income inequality.

Figure 1 – Scatterplot of Gini Coefficient vs. Mean Years of Schooling



Other variables include median age and government expenditure on age. These variables were used to strengthen the multiple linear regression models in an effort to prove that the link between education attained and income inequality is not as simple of a relationship that some may think. Government expenditure is measured by percentage of the country's GDP, to keep the analysis fair and consistent regardless of the size of the country. This helps in keeping the data steady across the entire study. This secondary variable was chosen due to Yang and Gao's surprising conclusions regarding government expenditure on education. With median age, it is also assumed there will be a moderately strong correlation between the median age and the Gini coefficient of a country. Age will provide a good representative of priority of education in a country, as education is commonly linked with positive health. Countries with a lower median age could also be linked with an inaccessibility to birth control, or a lack of education available to women. This could definitely factor into a country's income inequality.

Table 1 – Variable Descriptions

Variable	Description	Year	Units	Source
gini	Measure of income inequality, Gini Coefficient	2018	Percentage	Human Development Data Center
school	Mean years of schooling	2019	Years	Human Development Data Center
age	Median age of country	2020	Years	Human Development Data Center
govtexp	Govt. expenditure on education (of GDP)	2018	Percentage	Human Development Data Center
oecd	Whether a country is (1) or is not (0) an OECD country	2021	Dummy variable	OECD Organization

Table 2 – Variable Descriptive Statistics

Variable	Observations	Mean	St. Deviation	Min	Max
gini	147	38.31	7.888	24.2	63
school	147	8.595	3.241	1.6	14.2
age	147	29.89	9.473	15.2	48.4
govtexp	129	4.450	1.599	1	12.5
oecd	147	0.2245	0.4187	0	1

The source of the data is the Human Development Data Center. Fortunately, I was able to find all my necessary data from the same source. This keeps the sample size consistent across every single data set, so it limited the adjustments I needed to make. Unfortunately, the government expenditure variable was missing a few data points, but besides that, the number of observations was held constant at 147 for every variable. The data is also recent, most of it coming from 2018, 2019, and 2020. The Human Development Data Center is a trustworthy bank of information, as it uses the data from various continental agencies across the world who practice respectable surveying techniques.

1. MLR.1 Linear in Parameters: No variables were multiples of others. All of the regressions were run in Stata such that they are linear in parameters.

2. MLR.2 Random Sampling from Population: The data was sourced from the Human Development Data Center, where human development statistics were calculated by using various organizations across the world that practice random sampling from the population for every country included in this study. The sampling was random because the all data from the Human Development Data Center, as I did not pick and choose which countries to include. Data was observed in high- and low-income countries.
3. MLR.3 No Perfect Collinearity: Some variables are seen to have high correlation with one another, but the testing in Stata (Appendix C) proves there are no perfectly collinear variables. There were no exact linear relationships, so the regressions meet the third Gauss Markov assumption.
4. MLR.4 Zero Conditional Mean: Given any value of the explanatory values, the error term, u is always zero. The variables are assumed to be independent of the other variables in the system. We do not have any evidence to assure us this assumption is completely met, so we will proceed with caution.
5. MLR.5 Homoskedasticity: The error of the variance, u, is assumed to be relatively constant no matter the values of the regressors. Unfortunately, once again, this is very hard for us to verify, so we will proceed with caution.

Results

After checking to see if all of the assumptions are properly met, we can analyze the data using four different regression models. The STATA regression outputs for all four models can be found in Appendix B, while the STATA correlation outputs of each model can be found in Appendix C.

Model 1: Simple Regression Model

$$gini = \beta_0 + \beta_1 school + u$$

Firstly, a simple linear regression model is created to introduce the relationship between the Gini coefficient and average number of years of education completed.

Estimated Equation 1:

$$gini = 45.85 - 0.878school$$

This model's R-squared value is 0.13, meaning the correlation between average years of schooling and a country's Gini coefficient is weak. As expected by many, there is a negative sign in front of the coefficient on *school.* The negative coefficient of 0.878 means a 1-year increase in the average number of schooling for a country decreases its Gini coefficient by 0.878. This leads us to believe that an increase in the average years of schooling for a country will result in a decrease in the country's income inequality. The intercept of 45.85 does not reveal a great load of information to us, as the average years of schooling are not close to zero for a country. The number of observations being 147 paired with the weak R-squared value does lead us to start to believe the relationship between income inequality and education is not as correlated as one may intuitively think.

Model 2: First Multiple Regression:

$$gini = \beta_0 + \beta_1 school + \beta_2 oecd + \beta_3 govtexp + u$$

The first multiple regression model is created by adding in both of our secondary variables, government expenditure and OECD countries, into the original equation. Unfortunately, our sample size decreases from 147 to 129 due to our *govtexp* variable only having 129 data points.

Estimated Equation 2:

$$gini = 43.48 \ - 0.653 school - 3.618 oecd + 0.217 govtexp$$

This model's R-squared value is 0.1593, meaning the correlation between these explanatory variables and a country's Gini coefficient is weak. The addition of secondary explanatory variables leads us to have a more fitting regression model, but it is still a weak correlation. It is very intriguing to see the coefficient of *school* increase by about 0.2, suggesting that its impact on income inequality is becoming less strong with more variables taken into account. This is very interesting, as policymakers often attempt to fix the income inequality by expanding education. The small coefficient of 0.217 by *govtexp* also supports the idea that government expenditure on education expansion truly does not have the massive impact that it's usually hoped to have. The negative coefficient of -3.618 for *oecd* leads us to believe that whether or not a country is developed is the most impactful variable towards its Gini coefficient in this model. It also means that if a country is developed, their gini coefficient will decrease by 3.618, as the *oecd* variable is a dummy variable.

Model 3: Second Multiple Regression:

$$gini = \beta_0 + \beta_1 school + \beta_2 oecd + u$$

The second multiple regression model is created by removing government expenditure from the first multiple regression model. This way, the second multiple regression is consistent with every variable having 147 observations.

Estimated Equation 3:

$$gini = 44.53 - 0.63school - 3.602oecd$$

This equation's R-squared value is 0.1562, again showing a surprising, weak correlation. Upon removing *govtexp* this model was able to have a higher number of observations, 147. Once again, it is evident that whether or not a country is developed has a much bigger impact on a country's gini coefficient compared to the average years of schooling's impact on the gini coefficient. The correlation coefficient of school and oecd is one of the highest correlation coefficients (0.5318) of any of the relationships between the variables of every model. This intuitively makes sense, as average years of schooling probably is positively correlated with whether or not a country is developed. This second multiple regression is very similar to the first, reminding us that government expenditure on education is not very impactful.

Model 4: Third Multiple Regression:

$$gini = \beta_0 + \beta_1 school + \beta_2 age + \beta_3 govtexp + \beta_4 oecd + u$$

The third multiple regression model is created by adding in our last secondary variable, *age* into the last multiple regression model.

Estimated Equation 4:

$$gini = 48.57 + 0.232school - 0.432age + 0.154govtexp - 1.46oecd + u$$

This equation's R-squared value is 0.2391, again showing a weak correlation. Compared to the previous models, things have greatly changed. The variables *school*, *govtexp*, and *oecd* all are statistically insignificant, and they have very different coefficients when looking at the previous models. The coefficient of *school* is now positive which is very surprising, and the coefficient of *oecd* has been more than cut in half, decreasing its impact on the gini coefficient. The *govtexp* variable is now positive, but it still continues to have a miniscule impact on the model here. This model is very different compared to all the other ones, potentially leading us to believe that the median age of a country does not have a direct impact on inequality.

The table below gives a summary of the regression models above. The first number in each box represents the coefficient of the corresponding variable. The number in the parentheses is the standard deviation of the specified variable in the corresponding model, and the asterisks represent the statistical significance of the variable across all models.

Table 4: Regression Models Summary

Dependent Variable: Gini Coefficient

Ind. Variables	Model 1	Model 2	Model 3	Model 4
school	-0.878*** (0.189)	-0.653** (0.233)	-0.630*** (0.220)	0.232 (0.331)
oecd		-3.612* (1.86)	-3.602** (1.703)	-1.46 (1.87)
govtexp		0.217 (0.426)		0.154 (0.407)
age				-0.432*** (0.120)
Intercept	45.85*** (1.73)	43.48*** (2.475)	44.53*** (1.82)	48.57*** (2.75)
Observations	147	129	147	129
R-squared	0.13	0.156	0.159	0.239

*Significant at 10%, **5%, ***1%

Extensions

Robustness Test

In model 4, it was evident that the statistical significance of every single explanatory variable besides one, *age*, was insignificant at even the 10% level using a two-sided test. Looking at the correlation table for this model, these insignificant variables are moderately correlated, so it is now sensible to run a Robustness, or F-Test to see if these variables are jointly insignificant.

$$H_0: \hat{\beta}_1 = \hat{\beta}_3 = \hat{\beta}_4$$

$$H_1: H_0 \text{ is false}$$

From here, a restricted model was created from the original, unrestricted model.

Unrestricted Model:

$$gini = \beta_0 + \beta_1 school + \beta_2 age + \beta_3 govtexp + \beta_4 oecd + u$$

$$\text{n} = 129 \quad R^2 = 0.2391$$

Restricted Model:

$$gini = \beta_0 + \beta_2 age + u$$

$$\text{n} = 147 \quad R^2 = 0.2381$$

F-Test:

$$\text{F} = \frac{[(R_{ur}^2 - R_r^2)/q]}{[(1-R_{ur}^2)/(\text{n}-\text{k}-1)]}$$

Using the F-Test shown above, the F-statistic is 0.021, which is less than the F critical value of 3.00 for the 5% level. This signifies that the variables *school*, *govtexp*, and *oecd* are jointly insignificant at the 5% level for this model. This leads us to believe they can be removed from this model in an attempt to find a more fitting model for a country's income inequality.

Finally, we have the new model consisting of just the explanatory variable of *age* and its effect on a country's income inequality. As noted in the interpretation of model 4 in the Results section of this study, it is reasonable to assume that age does not have a direct relationship with a country's gini coefficient. Nonetheless, the STATA output for the simple regression model involving *gini* and *age* is given in Appendix B below.

Conclusions

This study was conducted with a hypothesis that mean years of schooling is a major factor when considering a country's income inequality. This hypothesis was somewhat supported by most of the regression models created in this study. It was evident throughout all the models that a country's mean years of schooling was negatively correlated with a country's gini coefficient, or, in other words, schooling decreases a country's income inequality. Obviously, this problem of income inequality is much more complex than just the relationship between schooling and inequality, evidenced by countries continuing to struggle with income inequality issues. This is shown to be true by all of the models having weak R^2 values. Multiple regression models were conducted with other explanatory variables with hopes of strengthening the model's goodness of fit and hopes of helping to identify which variables are most impactful towards this issue.

The added explanatory variables of *govtexp*, *oecd*, and *age* had interesting effects on the regression models. It seemed as though the R-squared value was always going to be weak through any course of action, probably a result of the fact that countries were used as observations in this study. This makes sense that there is no perfect fit for the relationship between a country's income inequality and the few explanatory variables we tested in this study, this is a testament to show how complex and different every country's scenario is. With the *govtexp* variable being generally insignificant throughout this study, it was intriguing to see if that was because of a multicollinear relationship with schooling. This was brought to my attention due to the fact that this variable's standard deviation was relatively large. Interestingly enough, the correlation coefficient between *school* and *govtexp* never reaches above 0.3, meaning that government expenditure on education is just simply not directly correlated with a country's income inequality.

This study shines a light on how complex the income inequality problem is throughout the entire world. It is very difficult to understand, and selecting a handful of things that directly contribute to income inequality is nearly impossible. I believe the best course of action would be to observe the countries that are consistently decreasing their income inequality gap year after year, and potentially learn from them to see what is beneficial. Singling out specific countries, cities, or states and comparing them to themselves could be a better way to study the income inequality problem, as their conditions stay mostly the same. It was brought to my attention through this study that it is so difficult to compare country to country because of the extremely different circumstances. It was definitely interesting diving into this intricate issue of national income inequality.

References

Alves, Nuno. "The Impact of Education on Household Income and Expenditure Inequality." *Taylor & Francis*, 27 Sept. 2011, www.tandfonline.com/doi/full/10.1080/13504851.2011.607125.

"Educational Attainment." *United States Census Bureau*, data.census.gov/cedsci/table?q=average+education&tid=ACSST1Y2019.S1501&hidePreview=false

Jerrim, John, and Lindsey Macmillan. "Income Inequality, Intergenerational Mobility, and the Great Gatsby Curve: Is Education the Key?" *OUP Academic*, Oxford University Press, 9 June 2015, academic.oup.com/sf/article/94/2/505/2583794#45335838.

"The OECD Member Countries." *OECD*, www.oecd.org/about/members-and-partners/.

Turčínková, Jana, and Jana Stávková. “Does the Attained Level of Education Affect the Income Situation of Households?” *Procedia - Social and Behavioral Sciences*, Elsevier, 10 Nov. 2012, www.sciencedirect.com/science/article/pii/S1877042812040578.

Yang, Juan, and Man Gao. “The Impact of Education Expansion on Wage Inequality.” *Taylor & Francis*, 7 Aug. 2017, www.tandfonline.com/doi/full/10.1080/00036846.2017.1361008?src=recsys.

Appendix A: Countries Used in Study

Albania	Greece	Palestine, State of
Algeria	Guatemala	Panama
Angola	Guinea	Papua New Guinea
Argentina	Guinea-Bissau	Paraguay
Armenia	Haiti	Peru
Australia	Honduras	Philippines
Austria	Hungary	Poland
Bangladesh	Iceland	Portugal
Belarus	India	Romania
Belgium	Indonesia	Russian Federation
Benin	Iran (Islamic Republic of)	Rwanda
Bhutan	Ireland	Saint Lucia
Botswana	Israel	Samoa
Brazil	Italy	Sao Tome and Principe
Bulgaria	Japan	Senegal
Burkina Faso	Jordan	Serbia
Burundi	Kazakhstan	Seychelles
Cabo Verde	Kenya	Sierra Leone
Cameroon	Korea (Republic of)	Slovakia
Canada	Kyrgyzstan	Slovenia

Central African Republic	Lao People's Democratic Republic	Solomon Islands
Chad	Latvia	South Africa
Chile	Lebanon	South Sudan
China	Lesotho	Spain
Colombia	Liberia	Sri Lanka
Comoros	Lithuania	Sudan
Congo	Luxembourg	Sweden
Congo (Democratic Republic of the)	Madagascar	Switzerland
Costa Rica	Malawi	Tajikistan
Croatia	Malaysia	Tanzania (United Republic of)
Cyprus	Maldives	Thailand
Czechia	Mali	Timor-Leste
Côte d'Ivoire	Malta	Togo
Denmark	Mauritania	Tonga
Djibouti	Mauritius	Tunisia
Dominican Republic	Mexico	Turkey
Ecuador	Micronesia (Federated States of)	Uganda
Egypt	Moldova (Republic of)	Ukraine
El Salvador	Mongolia	United Kingdom
Estonia	Montenegro	United States
Eswatini (Kingdom of)	Morocco	Uruguay
Ethiopia	Mozambique	Vanuatu
Fiji	Myanmar	Viet Nam
Finland	Namibia	Yemen
France	Nepal	Zambia
Gabon	Netherlands	Zimbabwe
Gambia	Nicaragua	
Georgia	Niger	
Germany	North Macedonia	
Ghana	Norway	
	Pakistan	

Appendix B: STATA Regression Outputs

```
regress gini school

      Source |       SS           df       MS      Number of obs   =       147
-------------+----------------------------------   F(1, 145)       =     21.67
       Model |  1181.23678         1  1181.23678   Prob > F        =    0.0000
    Residual |  7902.83314       145  54.5022975   R-squared       =    0.1300
-------------+----------------------------------   Adj R-squared   =    0.1240
       Total |  9084.06993       146   62.219657   Root MSE        =    7.3826

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      school |  -.8776529   .1885219    -4.66   0.000    -1.250259   -.5050471
       _cons |    45.8518    1.73102    26.49   0.000     42.43051    49.27309
------------------------------------------------------------------------------
```

```
. regress gini school govtexp oecd

      Source |       SS           df       MS      Number of obs   =       129
-------------+----------------------------------   F(3, 125)       =      7.90
       Model |  1273.81934         3  424.606445   Prob > F        =    0.0001
    Residual |  6720.04006       125  53.7603205   R-squared       =    0.1593
-------------+----------------------------------   Adj R-squared   =    0.1392
       Total |   7993.8594       128  62.4520266   Root MSE        =    7.3321

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      school |  -.6528646   .2325773    -2.81   0.006    -1.113164   -.1925652
     govtexp |   .2172027   .4261784     0.51   0.611    -.6262573    1.060663
        oecd |   -3.61826   1.860014    -1.95   0.054    -7.299458    .0629391
       _cons |   43.48236   2.475193    17.57   0.000     38.58364    48.38107
------------------------------------------------------------------------------
```

```
. regress gini school oecd

      Source |       SS           df       MS      Number of obs   =       147
-------------+----------------------------------   F(2, 144)       =     13.33
       Model |  1419.31207         2  709.656035   Prob > F        =    0.0000
    Residual |  7664.75786       144  53.2274851   R-squared       =    0.1562
-------------+----------------------------------   Adj R-squared   =    0.1445
       Total |  9084.06993       146   62.219657   Root MSE        =    7.2957

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      school |  -.6302129   .2199952    -2.86   0.005     -1.06505   -.1953759
        oecd |  -3.601615   1.702974    -2.11   0.036    -6.967672   -.2355584
       _cons |   44.53352   1.820682    24.46   0.000      40.9348    48.13223
------------------------------------------------------------------------------
```

```
. regress gini school age govtexp oecd

      Source |       SS           df       MS      Number of obs   =       129
-------------+----------------------------------   F(4, 124)       =      9.74
       Model |  1910.93259         4  477.733148   Prob > F        =    0.0000
    Residual |  6082.92681       124  49.0558613   R-squared       =    0.2391
-------------+----------------------------------   Adj R-squared   =    0.2145
       Total |   7993.8594       128  62.4520266   Root MSE        =     7.004

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      school |   .2321366   .3311569     0.70   0.485    -.4233157     .887589
         age |  -.4315576   .1197501    -3.60   0.000    -.6685766   -.1945386
     govtexp |   .1544788   .4074765     0.38   0.705    -.6520312    .9609889
        oecd |  -1.461249   1.874873    -0.78   0.437    -5.172147     2.24965
       _cons |   48.56947   2.753732    17.64   0.000     43.11906    54.01987
------------------------------------------------------------------------------
```

```
. regress gini age

      Source |       SS           df       MS      Number of obs   =       147
-------------+----------------------------------   F(1, 145)       =     45.30
       Model |  2162.49433         1  2162.49433   Prob > F        =    0.0000
    Residual |   6921.5756       145  47.7350041   R-squared       =    0.2381
-------------+----------------------------------   Adj R-squared   =    0.2328
       Total |  9084.06993       146   62.219657   Root MSE        =    6.9091

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.4062684   .0603607    -6.73   0.000    -.5255688    -.286968
       _cons |   50.44978   1.891788    26.67   0.000     46.71074    54.18883
------------------------------------------------------------------------------
```

Appendix C: STATA Correlation Outputs

```
. correlate gini school
(obs=147)

             |     gini   school
-------------+------------------
        gini |   1.0000
      school |  -0.3606   1.0000
```

```
. correlate gini school govtexp oecd
(obs=129)
```

	gini	school	govtexp	oecd
gini	1.0000			
school	-0.3657	1.0000		
govtexp	-0.0807	0.2550	1.0000	
oecd	-0.3259	0.5339	0.2851	1.0000

```
. correlate gini school oecd
(obs=147)
```

	gini	school	oecd
gini	1.0000		
school	-0.3606	1.0000	
oecd	-0.3289	0.5318	1.0000

```
. correlate gini school age govtexp oecd
(obs=129)
```

	gini	school	age	govtexp	oecd
gini	1.0000				
school	-0.3657	1.0000			
age	-0.4818	0.8219	1.0000		
govtexp	-0.0807	0.2550	0.2194	1.0000	
oecd	-0.3259	0.5339	0.5915	0.2851	1.0000

.